A1

a)

Anzahl der Regeln

|  |  |  |  |
| --- | --- | --- | --- |
|  | Contact Lenses | Cars | Kr vs Kp |
| Jrip | 2 Regeln | Einige Regeln mit guter generalisierung | 15 Regeln |
| Conjunctive | Nur default Rule | Nur default Rule | Eine Regel |
| Prism | Sehr viele Regeln (overfitted) | Sehr viele Regeln (of) | Sehr viele Regeln (of) |

Abbildung Anzahl der Regeln für die Jrip, Conjunctive und Prism Klassifizierer auf den Datensätzen Contact Lenses, Cars und KR vs KP

Bedingungen pro Regel

|  |  |  |  |
| --- | --- | --- | --- |
|  | Contact Lenses | Cars | Kr vs Kp |
| Jrip | 1-2 | 3-4 | 2-3 |
| Conjunctive | 0 | 0 | 1 |
| Prism | Ca 3 | Ca 5 | Ca 4 |

Abbildung Bedingungen pro Regel für die Jrip, Conjunctive und Prism Klassifizierer auf den Datensätzen Contact Lenses, Cars und KR vs KP

Vorhergesagte Klassen

|  |  |  |  |
| --- | --- | --- | --- |
|  | Contact Lenses | Cars | Kr vs Kp |
| Jrip | 3 Klassen 88% richtig | 4 Klassen 94 % richtig | 2 Klassen 99,5% richtig |
| Conjunctive | 3 Klassen 63% richtig | 4 Klassen 70% richtig | 2 Klassen 66% richtig |
| Prism | 3 Klassen 100% richtig | 4 Klassen 100% richtig | 2 Klassen 100% richtig |

Abbildung Vorhergesagte Klassen für die Jrip, Conjunctive und Prism Klassifizierer auf den Datensätzen Contact Lenses, Cars und KR vs KP

Conjunctive Rule: Versucht alles mit einer Regel abzudecken (allgemeinste lösung)

Prism: Versucht für jeden Datenpunkt eine Regel zu erstellen (overfitting)

Jrip: Mittelweg der eine generalisierte Lösung findet ohne zu viele Datenpunkte falsch zu klassifizieren und ohne zu viel overfitting zu erzeugen

b)

Default Rule:

|  |  |  |  |
| --- | --- | --- | --- |
|  | Contact Lenses | Cars | Kr vs Kp |
| Jrip | contact-lenses=none (12.0/0.0) | class=unacc (1164.0/7.0) | class=won (1667.0/6.0) |
| Conjunctive | contact-lenses = none | class = unacc | class = nowin |
| Prism | - | - | - |

Abbildung Default Rule für die Jrip, Conjunctive und Prism Klassifizierer auf den Datensätzen Contact Lenses, Cars und KR vs KP

Jrip:

Conjunctive: meistens wird die häufigste Klasse als default Rule ausgewählt ? / besser als nur eine Klassifikation

Prism: keine default Rule / keine Güte, da nicht abgedeckte Daten als unklassifiziert eingestuft werden

Jrip: meiste Attribute? Beste güte, da die Default Rule so gewählt wird, dass der fehler möglichst gering wird

c) Die Datenmengen mit den wenigsten Daten sind am einfachsten zu lernen (Contact Lenses) allerdings ergeben die Datensätze mit mehr Daten (kr vs kp) bessere Regeln.

d) Jrip ist allgemeiner, da hier nicht versucht wird Daten auswendig zu lernen (weniger overfitting), sondern durch eine generelle Regel auch mit nicht im Datensatz vorhandenen Daten umzugehen. Prism deckt alle Daten aus der Trainingsmenge mit Regeln ab und kann deshalb nicht mit unbekannten Daten umgehen, also nicht so gut dazulernen.

A2

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Abalone (2089) | Cross-val 5 | Cross-val 10 | Cross\_val 20 | Leave 1 | test | Validation |
| Correctly Classified | 19,87% | 19,63% | 18,48% | 19,00% | 22,31% | 18,73% |
| Incorrectly Class. | 80,13 % | 80,37 % | 81,52 % | 81,00 % | 77,69 % | 81,27 % |
| TP Rate | 0,199 | 0,196 | 0,185 | 0,19 | 0,223 | 0,187 |
| FP Rate | 0,147 | 0,149 | 0,153 | 0,15 | 0,15 | 0,154 |
| Precision | 0,206 | 0,182 | 0,186 | 0,181 | 0,229 | 0,147 |
| Recall | 0,199 | 0,196 | 0,185 | 0,19 | 0,223 | 0,187 |

Table : Vergleich verschiedener Kennzahlen (mittels 1x5 Cross-Validation, 1x10 Cross-Validation, 1x20 Cross-Validation, Leave-One-Out) und Bestimmung der Genauigkeit auf der Validierungsmenge für das Testset Abalone

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| balance\_scale (313) | Cross-val 5 | Cross-val 10 | Cross\_val 20 | Leave 1 | test | Validation |
| Correctly Classified | 78,59% | 77,32% | 76,36% | 79,87% | 81,79% | 80,45% |
| Incorrectly Class. | 21,41 % | 22,68 % | 23,64 % | 20,13 % | 18,21 % | 19,55 % |
| TP Rate | 0,786 | 0,773 | 0,764 | 0,799 | 0,818 | 0,804 |
| FP Rate | 0,182 | 0,193 | 0,201 | 0,172 | 0,155 | 0,168 |
| Precision | 0,723 | 0,712 | 0,703 | 0,735 | 0,758 | 0,743 |
| Recall | 0,786 | 0,773 | 0,764 | 0,799 | 0,818 | 0,804 |

Table Vergleich verschiedener Kennzahlen (mittels 1x5 Cross-Validation, 1x10 Cross-Validation, 1x20 Cross-Validation, Leave-One-Out) und Bestimmung der Genauigkeit auf der Validierungsmenge für das Testset

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Labor (29) | Cross-val 5 | Cross-val 10 | Cross\_val 20 | Leave 1 | test | Validation |
| Correctly Classified | 86,21% | 93,10% | 96,52% | 96,55% | 100,00% | 85,71% |
| Incorrectly Class. | 13,79 % | 6,90 % | 3,48 % | 3,45 % | 0,00 % | 14,29 % |
| TP Rate | 0,862 | 0,931 | 0,966 | 0,966 | 1 | 0,857 |
| FP Rate | 0,262 | 0,131 | 0,066 | 0,066 | 0 | 0,168 |
| Precision | 0,886 | 0,938 | 0,967 | 0,967 | 1 | 0,857 |
| Recall | 0,862 | 0,931 | 0,966 | 0,966 | 1 | 0,857 |

Table Vergleich verschiedener Kennzahlen (mittels 1x5 Cross-Validation, 1x10 Cross-Validation, 1x20 Cross-Validation, Leave-One-Out) und Bestimmung der Genauigkeit auf der Validierungsmenge für das Testset Labor

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Vehicle (423) | Cross-val 5 | Cross-val 10 | Cross\_val 20 | Leave 1 | test | Validation |
| Correctly Classified | 62,88% | 64,30% | 65,48% | 65,01% | 75,18% | 65,25% |
| Incorrectly Class. | 37,12 % | 35,70 % | 34,52 % | 34,99 % | 24,82 % | 34,75 % |
| TP Rate | 0,629 | 0,643 | 0,655 | 0,65 | 0,752 | 0,652 |
| FP Rate | 0,126 | 0,121 | 0,116 | 0,118 | 0,085 | 0,118 |
| Precision | 0,613 | 0,625 | 0,641 | 0,639 | 0,762 | 0,646 |
| Recall | 0,629 | 0,643 | 0,655 | 0,65 | 0,752 | 0,652 |

Table Vergleich verschiedener Kennzahlen (mittels 1x5 Cross-Validation, 1x10 Cross-Validation, 1x20 Cross-Validation, Leave-One-Out) und Bestimmung der Genauigkeit auf der Validierungsmenge für das Testset Vehicle

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Zoo (51) | Cross-val 5 | Cross-val 10 | Cross\_val 20 | Leave 1 | test | Validation |
| Correctly Classified | 84,31% | 84,31% | 88,24% | 86,27% | 94,12% | 84,00% |
| Incorrectly Class. | 15,69 % | 15,69 % | 11,76 % | 13,73 % | 5,88 % | 16,00% |
| TP Rate | 0,843 | 0,843 | 0,882 | 0,863 | 0,941 | 0,84 |
| FP Rate | 0,074 | 0,075 | 0,07 | 0,072 | 0,041 | 0,047 |
| Precision | 0,767 | 0,764 | 0,807 | 0,784 | 0,909 | 0,834 |
| Recall | 0,843 | 0,843 | 0,882 | 0,863 | 0,941 | 0,84 |

Table Vergleich verschiedener Kennzahlen (mittels 1x5 Cross-Validation, 1x10 Cross-Validation, 1x20 Cross-Validation, Leave-One-Out) und Bestimmung der Genauigkeit auf der Validierungsmenge für das Testset Zoo

1. Bei allen fünf Datensätzen kann man sehen, dass die Einschätzungen mit Croos-Validation einander ähnlich sind und im Gegensatz zur Einschätzung auf der Trainingsmengen generell schlechter ausfallen. Prinzipiell sollte man bei einer Genauigkeitsabchätzung den jeweils konservativsten Wert wählen. Dies ist bei Correctly Classified, TP Rate, Precision und Recall der niedrigste Wert und bei Incorrectly Classified, und FP Rate der höchste. Je nach Größe des Datensatzes sind die Ergebnisse der verschiedenen Abschätzungen unterschiedlich. Wie man sehen kann ist bei Abalone und balance\_scale die 1x20 Cross-Validation am negativsten und somit vermutlich die genaueste. Beide Datensätze sind eher größer. Bei den kleinen Datensätzen Labor und Zoo ist die 1x5 Cross-Validation am besten geeignet den Fehler abzuschätzen.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Abalone (2089) | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | Durchschnitt |
| Correctly Classified | 19,63% | 19,91% | 19,10% | 19,91% | 19,63% | 18,62% | 20,44% | 19,77% | 18,81% | 18,62% | 19,44% |
| Incorrectly Class. | 80,37 % | 80,09 % | 80,90 % | 80,09 % | 80,37 % | 81,38 % | 79,56 % | 80,23 % | 81,19 % | 81,38 % | 80,56% |
| TP Rate | 0,196 | 0,199 | 0,191 | 0,199 | 0,196 | 0,186 | 0,204 | 0,198 | 0,188 | 0,186 | 0,1943 |
| FP Rate | 0,149 | 0,148 | 0,151 | 0,149 | 0,149 | 0,153 | 0,147 | 0,149 | 0,152 | 0,152 | 0,1499 |
| Precision | 0,182 | 0,21 | 0,208 | 0,195 | 0,196 | 0,226 | 0,229 | 0,2 | 0,132 | 0,164 | 0,1942 |
| Recall | 0,196 | 0,199 | 0,191 | 0,199 | 0,196 | 0,186 | 0,204 | 0,198 | 0,188 | 0,186 | 0,1943 |

Table 10 mal 1x10 Cross-Validation mit verschiedenen Random-Seeds und der Durchschnitt der 10 Ergebnisse für die Trainingsmenge Abalone

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| balance\_scale (313) | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | Durchschnitt |
| Correctly Classified | 77,32% | 77,96% | 78,91% | 78,27% | 74,44% | 77,64% | 76,68% | 75,72% | 82,11% | 77,96% | 77,70% |
| Incorrectly Class. | 22,68 % | 22,04 % | 21,09 % | 21,73 % | 25,56 % | 22,36 % | 23,32 % | 24,28 % | 17,89 % | 22,04 % | 22,30% |
| TP Rate | 0,773 | 0,78 | 0,789 | 0,783 | 0,744 | 0,776 | 0,767 | 0,757 | 0,821 | 0,78 | 0,777 |
| FP Rate | 0,193 | 0,188 | 0,18 | 0,185 | 0,218 | 0,191 | 0,199 | 0,207 | 0,152 | 0,188 | 0,1901 |
| Precision | 0,712 | 0,718 | 0,726 | 0,72 | 0,685 | 0,714 | 0,706 | 0,698 | 0,756 | 0,717 | 0,7152 |
| Recall | 0,773 | 0,78 | 0,789 | 0,783 | 0,744 | 0,776 | 0,767 | 0,757 | 0,821 | 0,78 | 0,777 |

Table 10 mal 1x10 Cross-Validation mit verschiedenen Random-Seeds und der Durchschnitt der 10 Ergebnisse für die Trainingsmenge balance\_scale

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Labor (29) | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | Durchschnitt |
| Correctly Classified | 93,10% | 96,55% | 96,55% | 96,55% | 93,10% | 96,55% | 96,55% | 96,55% | 89,66% | 89,66% | 94,48% |
| Incorrectly Class. | 6,90 % | 3,45 % | 3,45 % | 3,45 % | 6,90 % | 3,45 % | 3,45 % | 3,45 % | 10,34 % | 10,34 % | 5,52% |
| TP Rate | 0,931 | 0,966 | 0,966 | 0,966 | 0,931 | 0,966 | 0,966 | 0,966 | 0,897 | 0,897 | 0,9452 |
| FP Rate | 0,131 | 0,066 | 0,066 | 0,066 | 0,131 | 0,066 | 0,066 | 0,066 | 0,197 | 0,197 | 0,1052 |
| Precision | 0,938 | 0,967 | 0,967 | 0,967 | 0,938 | 0,967 | 0,967 | 0,967 | 0,911 | 0,911 | 0,95 |
| Recall | 0,931 | 0,966 | 0,966 | 0,966 | 0,931 | 0,966 | 0,966 | 0,966 | 0,897 | 0,897 | 0,9452 |

Table 10 mal 1x10 Cross-Validation mit verschiedenen Random-Seeds und der Durchschnitt der 10 Ergebnisse für die Trainingsmenge Labor

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Vehicle (423) | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | Durchschnitt |
| Correctly Classified | 64,30% | 63,12% | 62,41% | 62,17% | 62,17% | 66,90% | 61,47% | 63,59% | 63,12% | 63,36% | 63,26% |
| Incorrectly Class. | 35,70 % | 36,88 % | 37,59 % | 37,83 % | 37,83 % | 33,10 % | 38,53 % | 36,41 % | 36,88 % | 36,64 % | 36,74% |
| TP Rate | 0,643 | 0,631 | 0,624 | 0,622 | 0,622 | 0,669 | 0,615 | 0,636 | 0,631 | 0,634 | 0,6327 |
| FP Rate | 0,121 | 0,125 | 0,127 | 0,128 | 0,127 | 0,112 | 0,13 | 0,123 | 0,125 | 0,123 | 0,1241 |
| Precision | 0,625 | 0,615 | 0,603 | 0,601 | 0,603 | 0,658 | 0,586 | 0,622 | 0,619 | 0,612 | 0,6144 |
| Recall | 0,643 | 0,631 | 0,624 | 0,622 | 0,622 | 0,669 | 0,615 | 0,636 | 0,631 | 0,634 | 0,6327 |

Table 10 mal 1x10 Cross-Validation mit verschiedenen Random-Seeds und der Durchschnitt der 10 Ergebnisse für die Trainingsmenge Vehicle

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Zoo (51) | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | Durchschnitt |
| Correctly Classified | 84,31% | 86,27% | 86,27% | 86,27% | 88,24% | 84,31% | 86,27% | 84,31% | 86,27% | 84,31% | 85,69% |
| Incorrectly Class. | 15,69 % | 13,73 % | 13,73 % | 13,73 % | 11,76 % | 15,69 % | 13,73 % | 15,69 % | 13,73 % | 15,69 % | 14,31% |
| TP Rate | 0,843 | 0,863 | 0,863 | 0,863 | 0,882 | 0,843 | 0,863 | 0,843 | 0,863 | 0,843 | 0,8569 |
| FP Rate | 0,075 | 0,072 | 0,072 | 0,072 | 0,071 | 0,074 | 0,072 | 0,074 | 0,072 | 0,086 | 0,074 |
| Precision | 0,764 | 0,797 | 0,797 | 0,784 | 0,806 | 0,767 | 0,784 | 0,767 | 0,797 | 0,766 | 0,7829 |
| Recall | 0,843 | 0,863 | 0,863 | 0,863 | 0,882 | 0,843 | 0,863 | 0,843 | 0,863 | 0,843 | 0,8569 |

Table 10 mal 1x10 Cross-Validation mit verschiedenen Random-Seeds und der Durchschnitt der 10 Ergebnisse für die Trainingsmenge Zoo

1. Die 1x10 Cross-Validation unterscheidet sich kaum von den Validierungen in Aufgabenteil a). Unserer Meinung nach führt eine geschickte Auswahl an Random-Seeds nicht unbedingt zu einer besseren Abschätzung, weil die Auswahl zufällig ist und man keinen direkten Einfluss auf die ausgewählten Daten hat.
2. In den Tabellen 1-5 kann man erkennen, dass der Validation-Wert in den meisten Fällen etwas über dem konservativsten Wert der verschiedenen Validierungsmethoden liegt. Der Durchschnitt der 1x10 Cross-Validations aus der b) ist im Vergleich zu dem Validation-Wert in der letzten Spalte der Tabellen 1-5 nicht schlecht, aber man findet immer eine Cross-Validation aus der a) die besser ist.

A3

ROC Kurven für colic.arff für NaiveBayes und J48

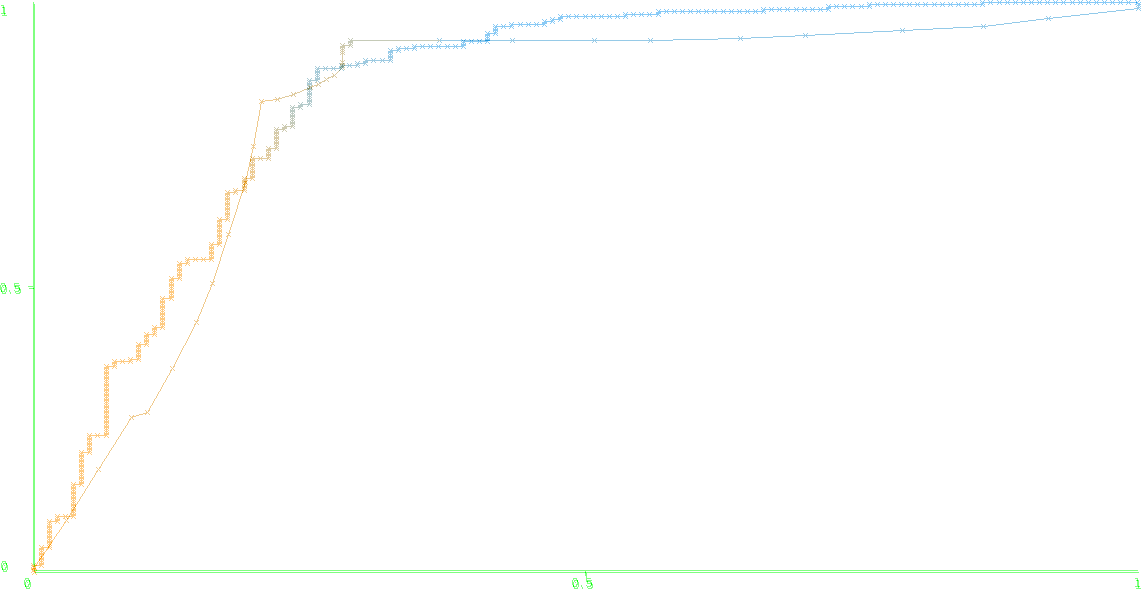


Abbildung ROC Kurven für den colic Datensatz mit NaiveBayes (obere) und J48(untere kurve) Klassifizierer.

1. - Die Fläche unter der NaiveBayes Kurve ist etwas größer als die unter der J48.

- NaiveBayes hat mehr Datenpunkte als J48.

- Die ROC Kurve von J48 liegt die meiste Zeit unter der des NaiveBayes Klassifizierers.

b) –Der NaiveBayes Algorithmus ist besser auf dem colic Datensatz, da die Fläche under der Kurve (AUC) größer ist und so der Anteil der True Positives gegenüber den False Positives höher ist, was bedeutet, dass die Daten besser Klassifiziert wurden.